DOCUMENT RESUME

ED 477 193 EA 032 529

AUTHOR	Conley, David T.; Brown, Richard S.
TITLE	Comparing State High School Assessments to Standards for Success in Entry-Level University Courses.
PUB DATE	2003-04-00
NOTE	30p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 21-25, 2003).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Standards; Accountability; Comparative Analysis; *Educational Assessment; Elementary Secondary Education; Evaluation; Evaluation Utilization; *Expectation; Higher Education; *Student Evaluation

ABSTRACT

This paper examines the relationship between high-school examinations and university expectations for a well-prepared student. Its purpose is to explore the alignment between state, standards-based assessment systems and the expectations students face once they undertake university studies. It uses a proven methodology for analyzing the congruence between the content of state high-school assessments and a set of standards keyed to the knowledge and skills necessary for success in select American research universities. An analysis of 30 state assessments in English/language arts (ELA) and 30 state assessments in mathematics revealed, on average, that more than half of the state ELA assessments were in alignment with university expectations. However, only one-third of the math assessments were shown to be in alignment. Overall, it appears that the alignment between state assessments and university expectations is better than expected in the areas of categorical concurrence, depth of knowledge, and balance of representation, especially given that state tests were not designed with postsecondary standards as reference points. However, when evaluated via a summary alignment index, it appears that most state assessments are not well aligned to the knowledge and skills for university success in mathematics, whereas most are so aligned in the area of ELA. (Contains 11 references.) (RJM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 477 193

Comparing State High School Assessments to Standards for Success in Entry-Level University Courses

David T. Conley, Ph.D.

Associate Professor, University of Oregon

Director, Center for Educational Policy Research

541-346-6155

conley@uoregon.edu

Richard S. Brown, Ph.D.

Assistant Professor

University of California, Irvine

Director,

Center for Research on Evaluation, Assessment, and Measurement

949-824-1561

rsbrown@uci.edu

Presented at the Annual Meeting of the American Educational Research Association

Chicago, IL, April 24, 2003

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

B. CONLEY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

EA032535

BEST COPY AVAILABLE

This study examines the relationship between state high school examinations and university expectations for a well-prepared student. A well-prepared student is defined as one who can succeed in entry-level courses in core academic areas within the general education portion of the university. The goal of the research is to determine if high school tests and university expectations are aligned and to create better linkages between high school reform efforts and postsecondary success. The purpose of the research is to ascertain the alignment between state standards-based assessment systems and the expectations students face once they undertake university studies.

The study is an outgrowth of emerging theories of systems coherence and alignment advanced by Fuhrman and others (Clune, 1998; Fuhrman, 1993, 2001; Goertz & et al., 1996) who assert that aligning local educational systems with state standards and assessments produces superior learning results. Signaling theory also frames this analysis. It states that high school students, teachers, administrators, and others receive signals from state standards and assessments and postsecondary admission requirements, among other sources, about what is important to teach and learn in high school. If the signals sent are unclear to those who receive them or cannot be easily linked to existing programs in ways that suggest how such programs should change, they tend to be misinterpreted or ignored.

Systems coherence and alignment are implicit assumptions underlying much of the standards-based policy adopted by states over the past decade. Signaling has not been identified as explicitly as a motivator for state legislators, but appears to be more important as states link assessments with accountability systems designed to maximize alignment and clarify signals. Postsecondary education, however, has not been a party to

this process for the most part. Although individual professors have participated in standard writing and assessment development, very few states have sought to connect their standards and assessments directly with postsecondary education admission requirements or even with broad notions of college readiness.

Whereas Weick's description of loosely coupled systems has been a very useful way to describe public educational governance for over two decades, current policies adopted by the state and federal government call into question whether education will continue as a loosely coupled system or whether the loose coupling will frustrate attempts at alignment. Examining the relationship between state high school assessments and college readiness standards helps to determine the current strength of the coupling between systems, based on tests, and to ascertain the implications for misalignment, if it exists.

This study applies a proven methodology for analyzing the congruence between the content of state high school assessments and a set of standards keyed to the knowledge and skills necessary for success in select American research universities. This study has much in common with content alignment work that has taken place comparing state test content to state and national content standards. It differs, however, in one important way. The state assessments in this study were not reviewed to determine their alignment with their own test specifications or with their state content standards. Rather, the items on these tests were analyzed to determine their alignment with the first set of standards that specifies the knowledge and skills that university faculty have identified as important for success in entry-level university courses (Conley, 2003).

Recently, Impara (2001) detailed several models for conducting content alignment studies, ranging from low to high levels of complexity. While studies of content alignment have been undertaken for many years, more concentrated effort has been devoted to creating new models for conducting such studies in recent years. Alignment methodology generally employs content experts to rate the match between a test item's content and the content standards referenced in test specifications. The rating scales employed capture information on the breadth of content coverage and the depth of content coverage (cognitive complexity). The level of agreement among raters is determined and presented along with the level of content match between the assessment and the content standards or test specification.

A variation of this approach was utilized by Le, Hamilton, & Robyn (2000) in their alignment among secondary and post-secondary assessments in California. This approach is certainly viable for the current context, however it is limited in the amount of information that it provides. A more moderately complex approach is described by LaMarca, et al. (2000). This approach enhances the simpler model by broadening the focus to include not only content match and depth of coverage, but also balance of coverage, performance match, and accessibility. Impara states, "because Webb's conceptualization of alignment is comprehensive and extends far beyond the two previously described models, it can be adapted for use in virtually any context" (p. 4). As a result, this approach, with a few modifications, is the procedure used for aligning state assessments with the KSUS for university success in the S4S project.

Like the Webb (1999) approach, the methodology employs in this study focuses on content analysis, specifically in the areas of categorical concurrence, depth of

knowledge, range of knowledge, and balance of representation. However, this research also develops a single index-value of alignment for a given test that consolidates alignment criteria data across multiple dimensions (i.e., range of knowledge, depth of knowledge, and balance of knowledge) and many standards within the KSUS criteria.

For example, in the area of mathematics, the alignment analysis using Webb's methodology generates four criterion values (categorical concurrence, range of knowledge, depth of knowledge, and balance) for each of six standards within the KSUS mathematics standards (Computation, Algebra, Geometry, Math Reasoning, Trigonometry, and Statistics). Thus, a total of twenty-four alignment measures for each state mathematics assessment, (sixteen for English/Language Arts exams), result from the Webb approach.

ALIGNMENT RATING PROCESS

Raters within each mathematics or English/Language Arts made three judgments. within their specific subject area. Specifically, 1) they rated the Depth of Knowledge for each KSUS standard (they did this one time only); 2) they rated the Depth of Knowledge for each state assessment item; and 3) they determined the match between individual KSUS standards and individual state assessment items.

The KSUS standards were the developed through a multi-step, multi-method process that began with nine meetings conducted across the nation at universities that are members of the Association of American Universities (AAU). Approximately 400 faculty members and administrators who teach or work with incoming students provided input on the key knowledge and skills needed for success in entry-level courses at the 20 institutions represented at the meetings. The project was sponsored by the following

universities: the Universities of California, Berkeley; Illinois; Iowa; Michigan; Minnesota; Nebraska; Wisconsin; Harvard University; Indiana University; New York University; Pennsylvania State University; Rutgers, the Massachusetts Institute of Technology; the University of Southern California, Rice University. The input received at these meetings was analyzed and synthesized. It was reviewed by an external consulting group with expertise in standard setting, Mid-Continent Research for Education and Learning, and finally by a Content Review Panel consisting of professors from AAU universities with particular expertise and content knowledge.

The raters first rated of the KSUS standards on a five-point scale adopted from Marzano (2001) to determine the depth of knowledge of each. The points of the scale were Retrieval, Comprehension, Analysis, Utilization, and Goal Setting/Monitoring. Marzano's scale is designed to reflect increasing cognitive complexity, more sophisticated uses of knowledge, and progressively higher levels of meta-cognitive functioning. The raters then used the five-point scale to assign a depth-of-knowledge rating to each assessment item. All rater within a discipline rated all items. The third rating task required raters to review each assessment item and determined whether it address any of the KSUS standards.

The result of these three analyses was a matrix (KSUS by assessment items) that identified which test item or items covered the content of a particular KSUS, how many KSUS were addressed by how many state test items, and which KSUS were not addressed by any state test items.

This process, known as content focus alignment (Webb, 1997, 1999), yields scales to which criterion values can be applied to reach a conclusion about how well each assessment met or exceeded an acceptable level of alignment. The four alignment criteria analyzed and the manner in which they are addressed were Categorical Concurrence, Depth of Knowledge, Range of Knowledge, and Balance of Representation.

Categorical Concurrence

The second analysis involved determining the match between the KSUS statements and the assessment items. The goal was to determine two things: First, do state assessments cover areas deemed important to university success; and, second, are the KSUS statements in their current format useful for determining such relationships? Raters were asked to identify which (if any) KSUS statements were addressed by each assessment item. These ratings were then summarized across raters and to determine the average number of items for each assessment addressing each standard within the KSUS. This criterion could range from 0 to the total number of items included on the examination. The recommended benchmark for this alignment criterion is that at least 6 items on average are aligned to each KSUS.

Depth of Knowledge Consistency

Based on the Depth of Knowledge ratings, test items were compared with corresponding KSUS statements to determine if state assessment items were at the same, higher, or lower level of cognitive complexity as the KSUS statement. Values for this criterion range from 0 to 1.0, with a recommended benchmark of >=.50. That is, for a state assessment to be considered adequately aligned to a given KSUS standard with respect to Depth of Knowledge, at least half of the assessment items addressing that

standard should be at or above the cognitive complexity level of the corresponding KSUS objective statement.

Range of Knowledge

For those items that did match with KSUS statements, the range of the match was determined by tallying the number of each KSUS objective statements that were addressed by one or more assessment item. This criterion gives an estimate of the breadth of expected knowledge addressed. As Webb states, "The range-of-knowledge criterion is used to judge whether a comparable span of knowledge expected of students by a standard is the same as, or corresponds to, the span of knowledge that students need in order to correctly answer the assessment items/activities" (Webb, 1999, p. 8). This criterion ranges in value from 0 to 1.0, with a recommended benchmark of >=.50. That is, it is suggested that for an assessment to be adequately aligned to a KSUS standard with respect to Range of Knowledge, at least half of the objectives within a KSUS standard should be addressed by at least one assessment item.

Balance of Representation

The final analysis identified the distribution of items matched with specific KSUS objectives. This criterion helps indicates the extent assessment items are evenly distributed across the KSUS objective that are addressed by each state assessment. Like the Depth of Knowledge and Range of Knowledge criteria, the Balance Index also ranges from 0 to 1.0, but has a recommended benchmark of .70 rather than .50 as in these other alignment measures.

DATA SOURCES

Thirty-five states have state-level high school examinations. Twenty-four of these currently have or will have implications for graduation. The rest are connected to state accountability systems. All states with high school examinations were invited to participate in the study. Assessments from 20 states were ultimately analyzed. The remaining state assessments were not analyzed for a number of reasons. Six states were in the process of modifying their current assessment or developing a new assessment to be implemented within the next year so that results of this analysis would have appeared about the time the new version of the assessment was being implemented. Every state had some form of math and English/Language Arts examination, although Illinois utilized a version of the ACT in English and math. In addition to multiple-choice tests, which every state had, 8 states had constructed response items included in their tests, and 4 had performance tasks, generally a writing prompt.

RESULTS AND CONCLUSIONS

Results are reported for each of the four alignment analysis methods, categorical concurrence, depth of knowledge, range, and balance.

Categorical Concurrence

State English/Language Arts assessments demonstrated what can be described as reasonable categorical concurrence with KSUS standards. On average, each KSUS standard had over 10 items addressing it for each of the state assessments (M =10.9, SD =- 17.1).

However, looking at the average number of items across various ELA standards disguises certain shortcomings. When viewed in terms the frequency with which a test meets the alignment benchmark adopted for this study of 6 or more items for a given

standard, only 42% of the assessments items met the mark. That is, more than half of the time the state high school assessment did not have enough items addressing the KSUS standard in English/Language Arts to be considered adequately aligned.

Examining alignment in terms of each of the four KSUS standard areas yields a more nuanced interpretation. Reading and Comprehension was adequately aligned 25 out of 30 times (83.3%) and writing half the time (15 of 30). The results were different for the KSUS standards of Critical Thinking and Research Skills. Only 30% of the assessments (9 of 30) met the mark for Critical Thinking and only 1 (3.3%) did so for Research Skills.

The analysis of mathematics assessments yielded similar results. The average across all standards was more than 11 items per assessment addressing each standard (M=11.6, SD=10.4). Once again, averages were quite different across the KSUS standards. While overall concurrence was higher than in English/Language Arts assessments (59.24% vs. 41.67%), two KSUS standards accounted for most of the alignment. Math assessments met the minimum 6 items per standard threshold in Computation on 90% of the tests and in Math Reasoning on 93.3% of the tests. The other KSUS math standards were less frequently aligned. State tests met the alignment criterion for the Algebra standard 63% of the time (19 of 30); with Geometry 60% of the time (18 of 30); and with the Trigonometry standard once among the 30 tests (3%). The KSUS standard of Statistics was not adequately addressed by any state assessment.

Depth of Knowledge Consistency

There appears to be greater alignment in terms of depth of knowledge for the ELA assessments than was seen for the categorical concurrence criterion. On average across

all standards, more than 60% of the assessment items were at or above the cognitive complexity level of the corresponding KSUS objective. With respect to meeting the recommended threshold for adequate alignment regarding depth of knowledge, across all standards more than 73% of assessments/standards comparisons resulted in at least half of the assessment items being at least as cognitively demanding as the corresponding objective. While these alignment accomplishments are not equivalent across standards, they are not as different with respect to this criterion as was seen with categorical concurrence. The state assessments were adequately aligned to the Reading and Comprehension standard the most frequently (93.3%), followed by Writing (70%), Research Skills (66.7%), then Critical Thinking (63.3). Thus, more frequently than not, all of the KSUS standards in English/Language Arts were addressed by state assessment with items of appropriate cognitive complexity.

In the area of Mathematics, even more alignment with respect to depth of knowledge was found. On average almost 70% of the assessment item/standard comparisons resulted in items deemed at or above the cognitive complexity of the corresponding KSUS objective. Further, over 82% (130/157) of the assessments met the threshold for adequate alignment regarding depth of knowledge across all standards. Depth of knowledge alignment was frequently established in almost all areas of mathematics, with only the standard of Trigonometry (36.7%) failing to exhibit high rates of adequate alignment. All other KSUS math standards were met with adequate depth of knowledge alignment over 80% of the time.

Range of Knowledge

The criterion for alignment regarding range of knowledge is more difficult to meet than the previous alignment criteria. For this criterion, each assessment should include items that address the majority of objectives within each KSUS standard. This task is made difficult by two issues. First, the assessments are limited in the number of items they can contain due to time and other testing constraints. Second, the KSUS objectives within each standard are fairly extensive and comprehensive. Thus, it is foreseeable that this alignment criterion will be met with less frequency by these state assessments than other alignment measures. In fact, this is exactly what we observed.

In Language Arts, on average, only 28.7% of the objectives within a KSUS standard were addressed and only about a quarter of the state tests (25.8%) met the threshold of addressing at least half of the KSUS objective within a given standard. This lack of alignment was fairly uniform across standard, with only Writing having as much as half (15/30) of the assessments address at least half of the KSUS objectives. The KSUS standard of Reading and Comprehension was adequately aligned with only 30% of the assessments and the standard of Critical Thinking was aligned only 23.3% of the time. Despite relatively few objectives for the KSUS standard of Research Skills, none of the 30 state assessments met the benchmark for adequately range of knowledge alignment for this standard.

Range of knowledge alignment was worse for mathematics. As in Language Arts, only about 28% of the objectives within the KSUS standards were addressed, but less than 13% of the assessments met the benchmark of addressing at least half of the objectives for a given standard. This lack of range of knowledge alignment was

consistent across standard. Only in the areas of Computation (23.3%) and Geometry (20.0%) did more than 1 in 5 state assessments address at least half of the objectives within the standard. Only 3 did so in Algebra (10%), 2 in Trigonometry (6.7%), and only one did so for Math Reasoning and Statistics.

Balance of Representation

While alignment values across the state assessments were lower relative to the range of knowledge criterion, the balance of representation alignment criterion yielded higher results. Most assessments addressing a given KSUS standard do so with a balanced set of items. In the area of English/Language Arts, the average balance index across all standards was .63 and in Mathematics it was .57. In ELA 79.1% of the assessments demonstrated adequate alignment across all standards relative to balance of representation. Similarly, 75.2% did so in the area of Mathematics.

Some standards were addressed in a more balanced fashion than others. In ELA, the KSUS standard of Reading and Comprehension was adequately aligned almost always (96.7%). The state assessments were well aligned in the areas of Writing (83.3%), Critical Thinking (73.3%), and Research Skills (63.3%) as well. It is worth reiterating that if categorical concurrence is low for a KSUS standard, the alignment statistic is less meaningful. In other words, of only a few items on a state test were found to correspond with one or more KSUS, alignment is less significant.

In mathematics, the balance of representation of state assessments with the KSUS standards of Computation, Algebra, Geometry, and Math Reasoning was high (90%). Only in the areas of Trigonometry (33.3%) and Statistics (0%) did the state assessments not demonstrate high balance of representation.

ALIGNMENT FINDINGS BY STATE PROFILE

While the preceding analyses provide summary information regarding various alignment indices across a collection of state assessments, it is also desirable to view each state's assessments independently. Thus, for each state assessment, an alignment profile can be created displaying the four alignment criteria (Categorical Concurrence, Range of Knowledge, Depth of Knowledge, and Balance) for each standard within a given subject area. This display provides valuable information regarding the degree of alignment and illustrates where particular strengths and weaknesses exist. See Figures 1 and 2 below for an example of a given state assessment alignment profile in both English/Language Arts and Mathematics.

Figure 1

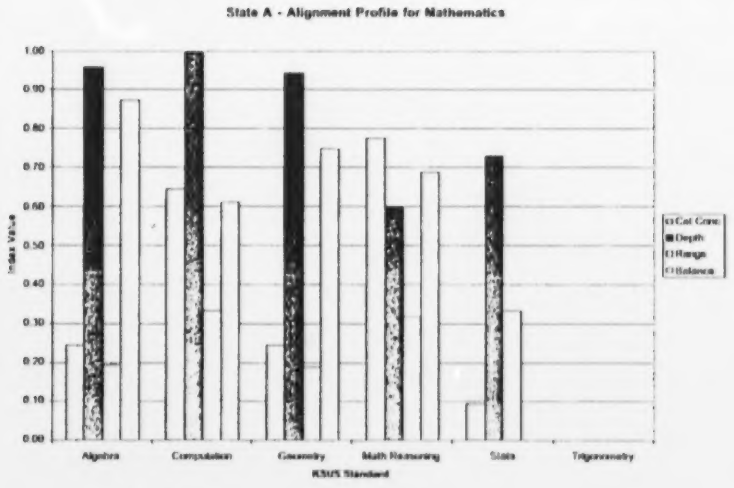
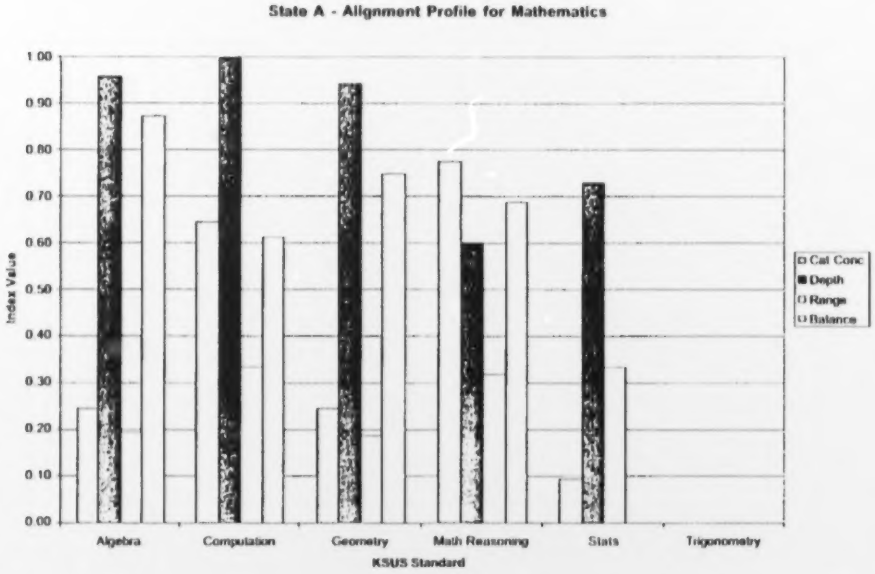


Figure 2

 BEST COPY AVAILABLE

In this English/Language Arts example, State A was more aligned with respect to Reading Comprehension and Writing than it was regarding the standards of Critical Thinking and Research Skills. Only in Reading Comprehension and Writing did this state have sufficient Range of Knowledge (>.50) as well as adequate Depth of Knowledge (>.50). For both Critical Thinking and Research Skills, this state assessment had lower alignment with regard to Range of Knowledge and Balance of Representation, while the Depth of Knowledge criterion value was high for the Critical Thinking standard.

Similarly, this state's math assessment showed depth of knowledge and balance of representation that met the criterion value, but range of knowledge did not for most



KSUS standard areas. This state assessment did not address the KSUS standard of Trigonometry at all.

CREATING A SUMMARY ALIGNMENT INDEX

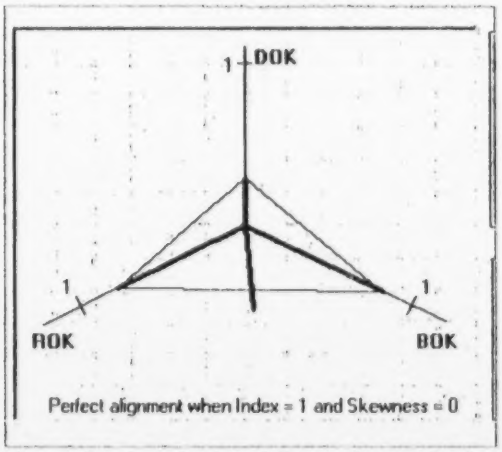
As the previous section points out, each state assessment generates a number of alignment index values across a collection of KSUS standards. While this information is valuable for identifying where alignment shortcomings or strengths may lie for a given assessment, the complexity of the information makes it quite difficult to draw overall conclusions about the alignment of each assessment to the KSUS standards. Further, comparisons between and among state assessments in a given subject area are very difficult.

To remedy this situation, a summary alignment index value (SAIV) was created for this study. The SAIV for a given subject area is the weighted average of the range of knowledge, depth of knowledge, and balance of representation measures.

For this study, categorical concurrence is not included in the summary calculations for a number of reasons. First, the scaling of the categorical concurrence index is not comparable to the other indices, which range from 0-1. Further, each of the other indices is independent across KSUS standard whereas the categorical concurrence measure is not entirely independent across standards. That is, unless assessment items addresses more than one KSUS standard, which would likely be inconsistent with typical test design objectives, the number of items addressing a given standard reduces the number of items available to address any other standard. Thus, if the categorical metric were to be reformulated to be the percent of items addressing a given standard rather than the number of items addressing a given standard, then the percent addressing Algebra, for instance, would limit the percent of items addressing the Computation standard. No such interdependency exists for the other alignment metrics across standards.

To create the SAIV, the alignment criterion values are analyzed to calculate the volume of the tetrahedron created by placing each of these values along one of three dimensions. Complete alignment (e.g., values of 1.0 on each of these three criteria) will yield a volume value of 1.0. Lesser values on each dimension will reduce the summary index value accordingly (see Figure 3 for a graphical representation of this idea).

Figure 3



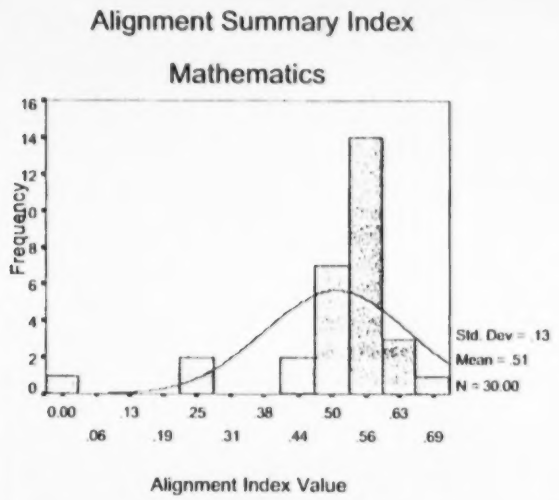
An alignment index value is calculated for each standard within a subject area. These values are then combined using a weighted average over the standards to create a single summary value for each state assessment based on the prevalence of KSUS objective statements per standard. These summary values (ranging from 0 to 1.0) are calculated for each state assessment and presented separately by subject area, thereby allowing for more interpretable comparisons across assessments.

DESCRIPTIVE STATISTICS OF ALIGNMENT INDEX

In terms of summary alignment, for this collection of 30 state assessments in English/language arts and 30 state assessments in mathematics, on average the alignment is comparable across subject area. The average summary alignment value in ELA is

BEST COPY AVAILABLE

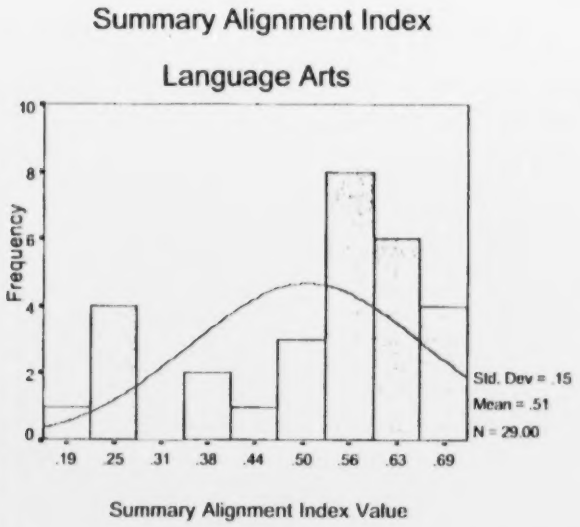
.507, and in mathematics it is .509 suggesting that the overall alignment across subject



areas is comparable for the average state assessment. See Figures 4-5 below for the distributions of each summary index by subject area.

Figure 4

Figure 5

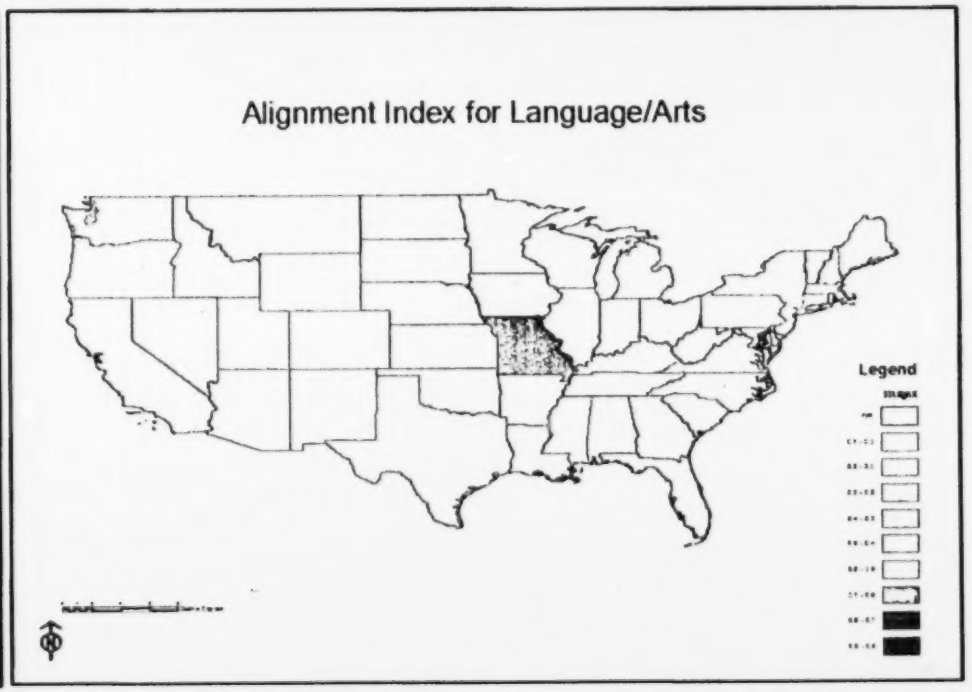


It appears that while the means and standard deviations are comparable, the skewness across the subject areas differs. If we used the recommended benchmark values for alignment suggested by Webb (.50 for depth of knowledge and range of knowledge, and .70 for balance of representation) to identify a benchmark for the summary measure, we would calculate a benchmark value of .5627 on the summary alignment index. That is, assessments yielding values below this point would be defined as not exhibiting adequate summary alignment, whereas those producing summary alignment values equal to or greater than this critical value would by definition be considered to exhibit adequate summary alignment.

Using these admittedly arbitrary benchmarks with this sample, one-third of the state math assessments demonstrate adequate summary alignment, and more than half (55.2%) of the state ELA assessments show adequate summary alignment. Thus, one may conclude that, *overall*, these state assessments were more frequently in alignment with the key knowledge and skills for university success in the area of English/language arts than they were in mathematics. To see the summary index values for each state's assessments, see Figures 6 and 7 below.

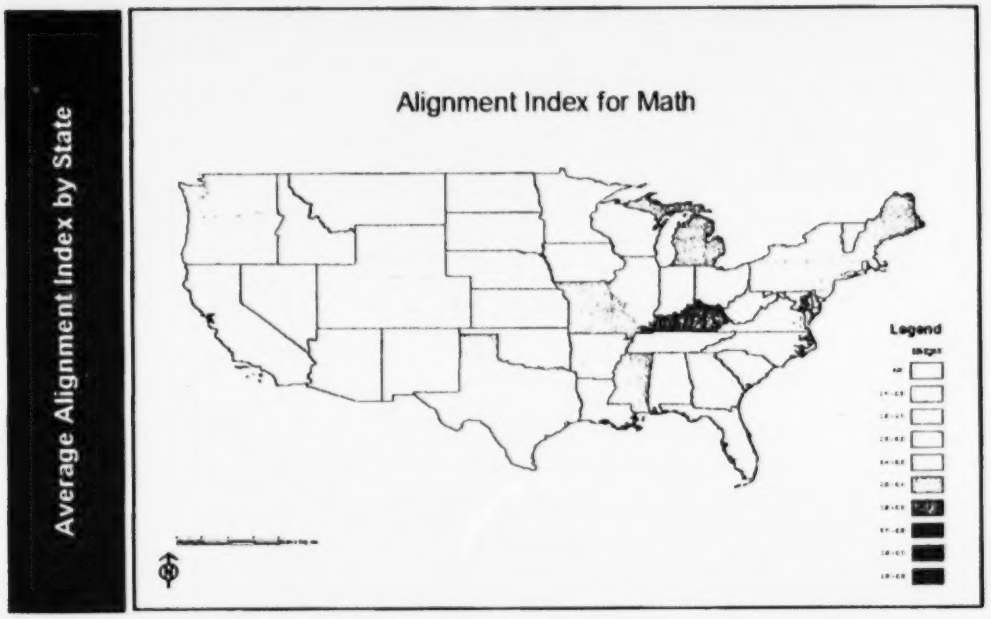
Figure 6

Average Alignment Index by State



BEST COPY AVAILABLE

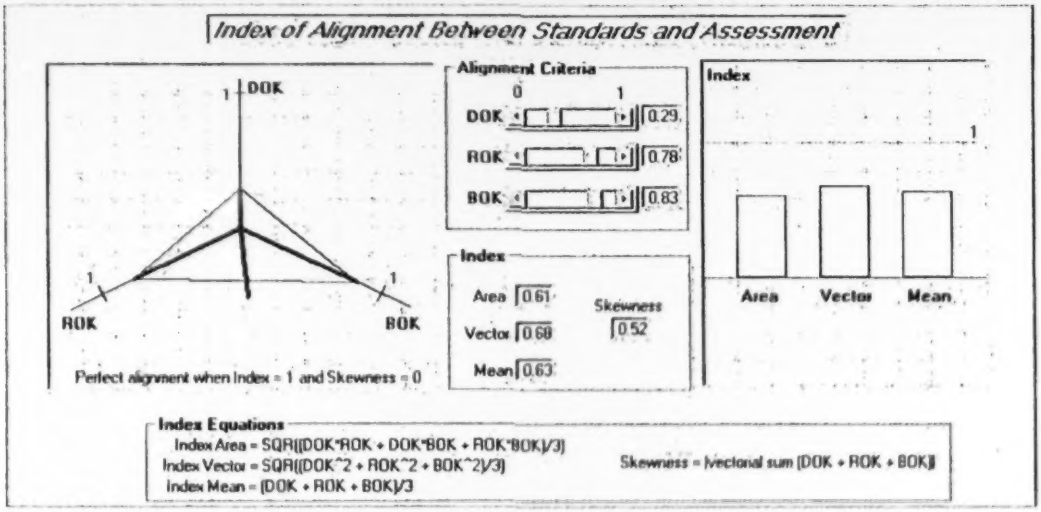
Figure 7



Other summary alignment measures were also explored, such as the simple average of the three alignment criterion values, the length of the vector from the origin to the center of the tetrahedron, as well as summary indices incorporating the categorical concurrence alignment criterion. Each of these approaches yielded quite comparable results. While a detailed discussion of these alternative summary alignment measures is not the focus of this paper, such investigations are the focus of our ongoing research efforts in this area. Figure 8 below provides a comparative example of three such measures.

BEST COPY AVAILABLE

Figure 8



SUMMARY OF FINDINGS

In general, the most significant findings can be summarized as follows:

- The alignment between state assessments and university KSUS statements is better than might have been expected in the areas of categorical concurrence, depth of knowledge, and balance of representation given that state tests were not designed with postsecondary standards as a reference point.
- Range of knowledge was better in math than in English, in part because there are more assessment items in math than in English.
- A frequent alignment shortcoming in state assessments stems from limited numbers of assessment items in the Critical Thinking and Research Skills areas for English and in the Trigonometry and Statistics area for mathematics exams, resulting in limited categorical concurrence and balance of representation. State tests do not attempt to

BEST COPY AVAILABLE

assess the potentially important knowledge areas in part because most students do not take the classes or master the necessary skills contained in these KSUS standards by the time the tests are given. Setting high school tests at the tenth grade generally precludes assessing trigonometry and statistics. Critical thinking and research skills can be tested with multiple-choice tests, but better measures exist, such as student research papers.

- For most states, the depth of knowledge consistency between KSUS statements and assessment items reaches sufficient levels of alignment. This is especially the case for almost all areas of mathematics, with Trigonometry being the sole exception. This can be explained at least in part based on the nature of the rating scale. Mathematics problems tend to fall into a more restricted range on the scale, and there are many KSUS standards in mathematics, including those that cover more basic math concepts. It becomes relatively easier to match problems with KSUS standards due at least in part to these factors. This is not automatically a design flaw in the KSUS or the rating system. It indicates that the potential for closer alignment between high school and college math courses may be a somewhat more manageable task and that state math assessments might be made to yield information useful to college admission and placement with some judicious changes.

- For most state assessments, range of knowledge alignment falls below acceptable benchmark levels. This is likely due to the fact that there are more KSUS objectives than assessment items in both mathematics and English. This finding suggests that as states revise their assessments (and the standards that underlie them), they may wish to examine more closely the mix of test items to determine if by adding or redistributing items, students might receive more useful diagnostic feedback about college readiness.

- For those KSUS areas that are addressed by state assessment items, there is often sufficient balance of representation to establish acceptable alignment on this criterion. This type of balance can be used to provide diagnostic feedback to students while they are still in high school to let them know how well prepared for college success they are in particular areas and what knowledge and skills they need to develop further if they plan to go to college. State tests do not currently provide this information. Admissions tests, such as the SAT and ACT, do so only in a limited fashion currently.

- When evaluated via a summary alignment index, it appears that most state assessments are not well aligned ***overall*** to the knowledge and skills for university success in the area of mathematics, but most are so aligned in the area of English/language arts. This is somewhat surprising given the specificity of the standards developed by the National Council of Teachers of Mathematics and the fact that these

standards have been broadly influential in the development of many state academic standards. Once again, at least a partial explanation is the number of mathematics KSUS standards. This, however, suggests that states may want to pay closer attention to the range of specific skills that are being tested on mathematics tests.

The study concludes that it is feasible to analyze state assessments in relation to key knowledge and skill statements and to generate valid results that could conceivably be of use to test designers, state education agency officials, and eventually to university administrators charged with making decisions about admissions and placement.

This study establishes a baseline or starting point against which predictive validity studies might be gauged. This information addresses issues of content validity of state tests for purposes of university admission and placement and helps establish the appropriate level of consequential validity that is appropriate for state examinations relative to college admissions.

Findings also suggest the limits of state testing, indicating that areas significant to university success receive little or no attention on such tests due to the limitations of their construction, time constraints, and the underlying state academic content standards.

Discussion

State high school standards and assessments are well established in many states as the primary policy framework for educational improvement. Federal legislation will require all states by 2004 to have a standards-based test in English and mathematics in grades 10-12. Science is added in 2005-2006. However, numerous states continue to experience problems with acceptance of their assessment systems by some educators,

parents, and policy makers, in part due to fears about the effects of the assessments on college preparation and college readiness.

High schools find themselves whip-sawn between state accountability requirements tied to test scores, upon which schools are judged, and demands to prepare students for college, which may require different courses and teaching methods. At the least, students are confronted with an increasing array of tests they must master, including SAT/ACT, Advanced Placement, SAT-II, International Baccalaureate, and now state assessments that can include multiple-choice tests along with end of course examinations. A policy that aligned state assessments with postsecondary expectations would help rationalize and perhaps simplify the assessment maze that is descending upon secondary schools.

A more fundamental issue is the content and methodology of the state assessments themselves. Are these the right formats for testing cumulative knowledge from twelve years of schooling? Even if they align in some ways with university expectations, do they really capture the more complex cognitive skills necessary to succeed in college? For example, the almost complete lack of actual writing in state assessments points out an obvious limitation of the formats adopted by states. Similarly, the passages students read in such tests come nowhere near the level of complexity and challenge students encounter from the day they enter the university.

End of course examinations fared better than did general achievement tests in terms of their alignment with university expectations. This suggests that the dozen or so states invested in such tests may have the potential to adapt these assessments to yield more useful data for university admission and to create better alignment with university

expectations. Such examinations also help alleviate the problem of testing all students against university standards. End of course exams allow greater alignment with postsecondary expectations to be incorporated as the course more closely approximates college preparation goals.

This line of research and its early findings helps raise questions of this nature in order to set the stage for more focused development work by states on how to make their assessments more valuable for a range of purposes that will cause students to want to do well on the assessments and that will engage teachers' energy and efforts to direct their efforts not toward test preparation, but toward the knowledge and skills that make a difference for future success.

References:

Clune, W. (1998). *Toward a Theory of Systemic Reform: The Case of Nine NSF Statewide Systemic Initiatives. Research Monograph No. 16.* Madison WI: National Institute for Science Education.

Fuhrman, S. H. (Ed.). (1993). *Designing Coherent Education Policy: Improving the System.* San Francisco: Jossey-Bass.

Fuhrman, S. H. (Ed.). (2001). *From the capitol to the classroom: Standards-based reform in the states.* Chicago: The University of Chicago Press.

Goertz, M. E., & et al. (1996). *The Bumpy Road to Education Reform. CPRE Policy Briefs*: Consortium for Policy Research in Education, Philadelphia, PA.

La Marca, P., Redfield, D., Winter, P., Bailey, A., & Despriet, L. (2000). *State standards and state assessment systems: A guide to alignment.* Washington, D.C.: Council of Chief State School Officers.

Le, V., Hamilton, L., & Robyn, A. (2000). Alignment among secondary and post-secondary assessments in California. In *Crucial Issues in California Education 2000.* Berkeley, CA: Policy Analysis for California Education.

Linstone, H. A., & Turoff, M. (1975). *The Delphi Method: Techniques and Applications.* Reading, MA: Addison-Wesley.

Marzano, R. J. (2001). *Designing a new taxonomy of educational objectives. Experts in assessment.* Thousand Oaks, California: Sage.

Webb, N. L. (1997). Research Monograph No. 6: Criteria for alignment of expectations and assessments in mathematics and science education. Washington, DC: Council of Chief State School Officers.

Webb, N. L. (1999). Research Monograph No. 18: Alignment of science and mathematics standards and assessments in four states. Madison, WI: National Institute for Science Education.

Webb, N. L. (2002, April). An analysis of the alignment between mathematics standards and assessments for three states. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

U.S. Department of Education
Institute of Education Sciences (IES)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

Reproduction Release

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title:	Comparing State High School Assessments to Standards for Success in Entry-Level University Courses
Author(s):	Conley, David T. & Brown, Richard S.
Corporate Source:	
Publication Date:	4-24-03

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign in the indicated space following.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
Level 1	Level 2A	Level 2B
☑	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g. electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche, or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature:		Printed Name/Position/Title: David T. Conley/Associate Professor & Director	
Organization/Address:	Center for Educational Policy Research 720 E. 13th Ave. Suite 201 Eugene, OR 97401	Telephone: 541-346-6153	Fax: 541-346-6154
		E-mail Address: conley@uoregon.edu	Date: 4-30-03

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:	
Address:	
Price:	

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:	
Address:	

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

Document Acquisitions Department
ERIC Clearinghouse on Educational Management
5207 University of Oregon
Eugene, Oregon 97403-5207